

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/568,168

Source: IFWP

Date Processed by STIC: 2/28/06

ENTERED



IFWP

RAW SEQUENCE LISTING

DATE: 02/28/2006

PATENT APPLICATION: US/10/568,168

TIME: 13:45:51

Input Set : A:\ITR0058P SEQLIST.TXT

Output Set: N:\CRF4\02282006\J568168.raw

```

4 <110> APPLICANT: Aurisicchio, Luigi
5     La Monica, Nicola
6     Giannetti, Patrizia
7     Ciliberto, Gennaro
9 <120> TITLE OF INVENTION: SYNTHETIC GENE ENCODING RHESUS MONKEY
10    CARCINOEMBRYONIC ANTIGEN AND USES THEREOF
13 <130> FILE REFERENCE: ITR0058P
C--> 15 <140> CURRENT APPLICATION NUMBER: US/10/568,168
C--> 15 <141> CURRENT FILING DATE: 2006-02-09
15 <150> PRIOR APPLICATION NUMBER: PCT/EP2004/009239
16 <151> PRIOR FILING DATE: 2004-08-17
18 <150> PRIOR APPLICATION NUMBER: 60/497,201
19 <151> PRIOR FILING DATE: 2003-08-22
21 <160> NUMBER OF SEQ ID NOS: 3
23 <170> SOFTWARE: FastSEQ for Windows Version 4.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 2118
27 <212> TYPE: DNA
28 <213> ORGANISM: Artificial Sequence
30 <220> FEATURE:
31 <223> OTHER INFORMATION: codon-optimized rhesus monkey CEA
33 <400> SEQUENCE: 1
34 atgggcagcc ccagcgcccc cctgcaccgc tgggtgcatcc cctggcagac cctgctgctg 60
35 accgccagcc tgctgacctt ctggaacccc cccaccaccg cccagctgac catcgagagc 120
36 cgccccttca acgtggccga gggcaaggag gtgctgctgc tggcccacaa cgtgagccag 180
37 aacctgttcg gctacatctg gtacaagggc gagcgcgtgg acgccagccg ccgcatcggc 240
38 agctgcgtga tccgcaccca gcagatcacc cccggccccg cccacagcgg ccgcgagacc 300
39 atcgacttca acgccagcct gctgatccac aacgtgaccc agagcgacac cggcagctac 360
40 accatccagg tgatcaagga ggacctggtg aacgaggagg ccaccggcca gttccgcgtg 420
41 tacccecgagc tgcccaagcc ctacatcagc agcaacaaca gcaaccccgt ggaggacaag 480
42 gacgccgtgg ccctgacctg cgagccccgag acccaggaca ccacctacct gtggtgggtg 540
43 aacaaccaga gctgccccgt gagccccgcg ctggagctga gcagcgacaa ccgcaccctg 600
44 accgtgttca acatcccccg caacgacacc accagctaca agtgcgagac ccagaacccc 660
45 gtgagcgtgc gccgcagcga ccccgtagcc ctgaacgtgc tgtacggccc cgacgcccc 720
46 accatcagcc ccctgaacac cccctaccgc gccggcgaga acctgaacct gacctgccac 780
47 gccgccagca accccaccgc ccagtacttc tggttcgtga acggcacctt ccagcagagc 840
48 acccaggagc tgttcatccc caacatcacc gtgaacaaca gcggcagcta catgtgccag 900
49 gccacaaca gcgccaccgg cctgaaccgc accaccgtga ccgccatcac cgtgtacgcc 960
50 gagctgccc agccctacat caccagcaac aacagcaacc ccacgagga caaggacgcc 1020
51 gtgacctga cctgcgagcc cgagaccagc gacaccacct acctgtggtg ggtgaacaac 1080
52 cagagcctga gcgtgagcag ccgcctggag ctgagcaacg acaaccgcac cctgaccgtg 1140
53 ttcaacatcc cccgcaacga caccaccttc tacgagtgcg agaccagaa ccccgtagc 1200
54 gtgcgccga gcgaccccg gaccctgaac gtgctgtacg gccccgacgc cccaccatc 1260

```

RAW SEQUENCE LISTING

DATE: 02/28/2006

PATENT APPLICATION: US/10/568,168

TIME: 13:45:51

Input Set : A:\ITR0058P SEQLIST.TXT

Output Set: N:\CRF4\02282006\J568168.raw

```

55 agccccctga acaccccccta ccgcgccggc gagaacctga acctgagctg ccacgccgcc 1320
56 agcaacccccg ccgcccagta cagctgggtc gtgaacggca ccttccagca gagcaccag 1380
57 gagctgttca tccccaacat caccgtgaac aacagcggca gctacatgtg ccaggccac 1440
58 aacagcggca ccggcctgaa ccgcaccacc gtgaccgcca tcaccgtgta cgtggagctg 1500
59 cccaagccct acatcagcag caacaacagc aaccccatcg aggacaagga cgccgtgacc 1560
60 ctgacctgcg agcccggtgc cgagaacacc acctacctgt ggtgggtgaa caaccagagc 1620
61 ctgagcgtga gccccgcct gcagctgagc aacggcaacc gcacacctgac cctgctgagc 1680
62 gtgaccgca acgacaccgg cccctacgag tgcggcatcc agaacagcga gagcgccaag 1740
63 cgcagcgacc ccgtgaccct gaacgtgacc tacggccccg acaccccat catcagcccc 1800
64 cccgacctga gctaccgag cggcgccaac ctgaacctga gctgccacag cgacagcaac 1860
65 cccagcccc agtacagctg gctgatcaac ggacacctgc gccagcacac ccagggtgctg 1920
66 ttcacagca agatcaccag caacaacagc ggcgctacg cctgcttcgt gagcaacctg 1980
67 gccaccggcc gcaacaacag catcgtgaag aacatcagcg tgagcagcgg cgacagcgcc 2040
68 cccggcagca gcggcctgag cgccgcgcc accgtgggca tcatcatcgg catgctggtg 2100
69 ggcgtggccc tgatgtga                                     2118

```

71 <210> SEQ ID NO: 2

72 <211> LENGTH: 705

73 <212> TYPE: PRT

74 <213> ORGANISM: Macaca Mulatta

76 <400> SEQUENCE: 2

```

77 Met Gly Ser Pro Ser Ala Pro Leu His Arg Trp Cys Ile Pro Trp Gln
78 1 5 10 15
79 Thr Leu Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro Pro Thr
80 20 25 30
81 Thr Ala Gln Leu Thr Ile Glu Ser Arg Pro Phe Asn Val Ala Glu Gly
82 35 40 45
83 Lys Glu Val Leu Leu Leu Ala His Asn Val Ser Gln Asn Leu Phe Gly
84 50 55 60
85 Tyr Ile Trp Tyr Lys Gly Glu Arg Val Asp Ala Ser Arg Arg Ile Gly
86 65 70 75 80
87 Ser Cys Val Ile Arg Thr Gln Gln Ile Thr Pro Gly Pro Ala His Ser
88 85 90 95
89 Gly Arg Glu Thr Ile Asp Phe Asn Ala Ser Leu Leu Ile His Asn Val
90 100 105 110
91 Thr Gln Ser Asp Thr Gly Ser Tyr Thr Ile Gln Val Ile Lys Glu Asp
92 115 120 125
93 Leu Val Asn Glu Glu Ala Thr Gly Gln Phe Arg Val Tyr Pro Glu Leu
94 130 135 140
95 Pro Lys Pro Tyr Ile Ser Ser Asn Asn Ser Asn Pro Val Glu Asp Lys
96 145 150 155 160
97 Asp Ala Val Ala Leu Thr Cys Glu Pro Glu Thr Gln Asp Thr Thr Tyr
98 165 170 175
99 Leu Trp Trp Val Asn Asn Gln Ser Leu Pro Val Ser Pro Arg Leu Glu
100 180 185 190
101 Leu Ser Ser Asp Asn Arg Thr Leu Thr Val Phe Asn Ile Pro Arg Asn
102 195 200 205
103 Asp Thr Thr Ser Tyr Lys Cys Glu Thr Gln Asn Pro Val Ser Val Arg
104 210 215 220
105 Arg Ser Asp Pro Val Thr Leu Asn Val Leu Tyr Gly Pro Asp Ala Pro

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/568,168

DATE: 02/28/2006

TIME: 13:45:51

Input Set : A:\ITR0058P SEQLIST.TXT

Output Set: N:\CRF4\02282006\J568168.raw

```

106 225                230                235                240
107 Thr Ile Ser Pro Leu Asn Thr Pro Tyr Arg Ala Gly Glu Asn Leu Asn
108                245                250                255
109 Leu Thr Cys His Ala Ala Ser Asn Pro Thr Ala Gln Tyr Phe Trp Phe
110                260                265                270
111 Val Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu Phe Ile Pro Asn
112                275                280                285
113 Ile Thr Val Asn Asn Ser Gly Ser Tyr Met Cys Gln Ala His Asn Ser
114                290                295                300
115 Ala Thr Gly Leu Asn Arg Thr Thr Val Thr Ala Ile Thr Val Tyr Ala
116 305                310                315                320
117 Glu Leu Pro Lys Pro Tyr Ile Thr Ser Asn Asn Ser Asn Pro Ile Glu
118                325                330                335
119 Asp Lys Asp Ala Val Thr Leu Thr Cys Glu Pro Glu Thr Gln Asp Thr
120                340                345                350
121 Thr Tyr Leu Trp Trp Val Asn Asn Gln Ser Leu Ser Val Ser Ser Arg
122                355                360                365
123 Leu Glu Leu Ser Asn Asp Asn Arg Thr Leu Thr Val Phe Asn Ile Pro
124                370                375                380
125 Arg Asn Asp Thr Thr Phe Tyr Glu Cys Glu Thr Gln Asn Pro Val Ser
126 385                390                395                400
127 Val Arg Arg Ser Asp Pro Val Thr Leu Asn Val Leu Tyr Gly Pro Asp
128                405                410                415
129 Ala Pro Thr Ile Ser Pro Leu Asn Thr Pro Tyr Arg Ala Gly Glu Asn
130                420                425                430
131 Leu Asn Leu Ser Cys His Ala Ala Ser Asn Pro Ala Ala Gln Tyr Ser
132                435                440                445
133 Trp Phe Val Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu Phe Ile
134                450                455                460
135 Pro Asn Ile Thr Val Asn Asn Ser Gly Ser Tyr Met Cys Gln Ala His
136 465                470                475                480
137 Asn Ser Ala Thr Gly Leu Asn Arg Thr Thr Val Thr Ala Ile Thr Val
138                485                490                495
139 Tyr Val Glu Leu Pro Lys Pro Tyr Ile Ser Ser Asn Asn Ser Asn Pro
140                500                505                510
141 Ile Glu Asp Lys Asp Ala Val Thr Leu Thr Cys Glu Pro Val Ala Glu
142                515                520                525
143 Asn Thr Thr Tyr Leu Trp Trp Val Asn Asn Gln Ser Leu Ser Val Ser
144                530                535                540
145 Pro Arg Leu Gln Leu Ser Asn Gly Asn Arg Ile Leu Thr Leu Leu Ser
146 545                550                555                560
147 Val Thr Arg Asn Asp Thr Gly Pro Tyr Glu Cys Gly Ile Gln Asn Ser
148                565                570                575
149 Glu Ser Ala Lys Arg Ser Asp Pro Val Thr Leu Asn Val Thr Tyr Gly
150                580                585                590
151 Pro Asp Thr Pro Ile Ile Ser Pro Pro Asp Leu Ser Tyr Arg Ser Gly
152                595                600                605
153 Ala Asn Leu Asn Leu Ser Cys His Ser Asp Ser Asn Pro Ser Pro Gln
154                610                615                620

```

RAW SEQUENCE LISTING

DATE: 02/28/2006

PATENT APPLICATION: US/10/568,168

TIME: 13:45:51

Input Set : A:\ITR0058P SEQLIST.TXT

Output Set: N:\CRF4\02282006\J568168.raw

```

155 Tyr Ser Trp Leu Ile Asn Gly Thr Leu Arg Gln His Thr Gln Val Leu
156 625 630 635 640
157 Phe Ile Ser Lys Ile Thr Ser Asn Asn Ser Gly Ala Tyr Ala Cys Phe
158 645 650 655
159 Val Ser Asn Leu Ala Thr Gly Arg Asn Asn Ser Ile Val Lys Asn Ile
160 660 665 670
161 Ser Val Ser Ser Gly Asp Ser Ala Pro Gly Ser Ser Gly Leu Ser Ala
162 675 680 685
163 Arg Ala Thr Val Gly Ile Ile Ile Gly Met Leu Val Gly Val Ala Leu
164 690 695 700
165 Met
166 705
169 <210> SEQ ID NO: 3
170 <211> LENGTH: 705
171 <212> TYPE: PRT
172 <213> ORGANISM: Macaca Mulatta
174 <400> SEQUENCE: 3
175 Met Gly Ser Pro Ser Ala Pro Leu His Arg Trp Cys Ile Pro Trp Gln
176 1 5 10 15
177 Thr Leu Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro Pro Thr
178 20 25 30
179 Thr Ala Gln Leu Thr Ile Glu Ser Arg Pro Phe Asn Val Ala Glu Gly
180 35 40 45
181 Lys Glu Val Leu Leu Leu Ala His Asn Val Ser Gln Asn Leu Phe Gly
182 50 55 60
183 Tyr Ile Trp Tyr Lys Gly Glu Arg Val Asp Ala Ser Arg Arg Ile Gly
184 65 70 75 80
185 Ser Cys Val Ile Arg Thr Gln Gln Ile Thr Pro Gly Pro Ala His Ser
186 85 90 95
187 Gly Arg Glu Thr Ile Asp Phe Asn Ala Ser Leu Leu Ile His Asn Val
188 100 105 110
189 Thr Gln Ser Asp Thr Gly Ser Tyr Thr Ile Gln Val Ile Lys Glu Asp
190 115 120 125
191 Leu Val Asn Glu Glu Ala Thr Gly Gln Phe Arg Val Tyr Pro Glu Leu
192 130 135 140
193 Pro Lys Pro Tyr Ile Ser Ser Asn Asn Ser Asn Pro Val Glu Asp Lys
194 145 150 155 160
195 Asp Ala Val Ala Leu Thr Cys Glu Pro Glu Thr Gln Asp Thr Thr Tyr
196 165 170 175
197 Leu Trp Trp Val Asn Asn Gln Ser Leu Pro Val Ser Pro Arg Leu Glu
198 180 185 190
199 Leu Ser Ser Asp Asn Arg Thr Leu Thr Val Phe Asn Ile Pro Arg Asn
200 195 200 205
201 Asp Thr Thr Ser Tyr Lys Cys Glu Thr Gln Asn Pro Val Ser Val Arg
202 210 215 220
203 Arg Ser Asp Pro Val Thr Leu Asn Val Leu Tyr Gly Pro Asp Ala Pro
204 225 230 235 240
205 Thr Ile Ser Pro Leu Asn Thr Pro Tyr Arg Ala Gly Glu Asn Leu Asn
206 245 250 255

```

RAW SEQUENCE LISTING

DATE: 02/28/2006

PATENT APPLICATION: US/10/568,168

TIME: 13:45:51

Input Set : A:\ITR0058P SEQLIST.TXT

Output Set: N:\CRF4\02282006\J568168.raw

```

207 Leu Thr Cys His Ala Ala Ser Asn Pro Thr Ala Gln Tyr Phe Trp Phe
208      260      265      270
209 Val Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu Phe Ile Pro Asn
210      275      280      285
211 Ile Thr Val Asn Asn Ser Gly Ser Tyr Met Cys Gln Ala His Asn Ser
212      290      295      300
213 Ala Thr Gly Leu Asn Arg Thr Thr Val Thr Ala Ile Thr Val Tyr Ala
214 305      310      315      320
215 Glu Leu Pro Lys Pro Tyr Ile Thr Ser Asn Asn Ser Asn Pro Ile Glu
216      325      330      335
217 Asp Lys Asp Ala Val Thr Leu Thr Cys Glu Pro Glu Thr Gln Asp Thr
218      340      345      350
219 Thr Tyr Leu Trp Trp Val Asn Asn Gln Ser Leu Ser Val Ser Ser Arg
220      355      360      365
221 Leu Glu Leu Ser Asn Asp Asn Arg Thr Leu Thr Val Phe Asn Ile Pro
222      370      375      380
223 Arg Asn Asp Thr Thr Phe Tyr Glu Cys Glu Thr Gln Asn Pro Val Ser
224 385      390      395      400
225 Val Arg Arg Ser Asp Pro Val Thr Leu Asn Val Leu Tyr Gly Pro Asp
226      405      410      415
227 Ala Pro Thr Ile Ser Pro Leu Asn Thr Pro Tyr Arg Ala Gly Glu Asn
228      420      425      430
229 Leu Asn Leu Ser Cys His Ala Ala Ser Asn Pro Ala Ala Gln Tyr Phe
230      435      440      445
231 Trp Phe Val Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu Phe Ile
232      450      455      460
233 Pro Asn Ile Thr Val Asn Asn Ser Gly Ser Tyr Met Cys Gln Ala His
234 465      470      475      480
235 Asn Ser Ala Thr Gly Leu Asn Arg Thr Thr Val Thr Ala Ile Thr Val
236      485      490      495
237 Tyr Val Glu Leu Pro Lys Pro Tyr Ile Ser Ser Asn Asn Ser Asn Pro
238      500      505      510
239 Ile Glu Asp Lys Asp Ala Val Thr Leu Thr Cys Glu Pro Val Ala Glu
240      515      520      525
241 Asn Thr Thr Tyr Leu Trp Trp Val Asn Asn Gln Ser Leu Ser Val Ser
242      530      535      540
243 Pro Arg Leu Gln Leu Ser Asn Gly Asn Arg Ile Leu Thr Leu Leu Ser
244 545      550      555      560
245 Val Thr Arg Asn Asp Thr Gly Pro Tyr Glu Cys Gly Ile Gln Asn Ser
246      565      570      575
247 Glu Ser Ala Lys Arg Ser Asp Pro Val Thr Leu Asn Val Thr Tyr Gly
248      580      585      590
249 Pro Asp Thr Pro Ile Ile Ser Pro Pro Asp Leu Ser Tyr Arg Ser Gly
250      595      600      605
251 Ala Asn Leu Asn Leu Ser Cys His Ser Asp Ser Asn Pro Ser Pro Gln
252      610      615      620
253 Tyr Ser Trp Leu Ile Asn Gly Thr Leu Arg Gln His Thr Gln Val Leu
254 625      630      635      640
255 Phe Ile Ser Lys Ile Thr Ser Asn Asn Asn Gly Ala Tyr Ala Cys Phe

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/568,168

DATE: 02/28/2006

TIME: 13:45:52

Input Set : A:\ITR0058P SEQLIST.TXT

Output Set: N:\CRF4\02282006\J568168.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No

L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date